Blockchain-enabled RF Radiation Exposure Level Measurements in Wireless Mobile Networks

Final Project Report

Submitted by

Dr. Geeth Priyankara
Dr. Chathura Seneviratne
Dr. Madhusanka Liyanage
Inushi Gunatilake A.A.P.I.
Nivin Madawalage
Tharani Thalgahagoda

Department of Electrical and Information Engineering
Faculty of Engineering
University of Ruhuna
Sri Lanka

March 2024

Abstract

The expansion of telecommunication providers and their efforts to expand network coverage in Sri Lanka has led to an alarming increase in Radio Frequency (RF) radiation exposure. This heightened exposure has raised concerns about its potential adverse effects on human health, including infertility, cancer, crib death, and DNA damage, particularly in fetuses and infants. Despite the country's growing number of cancer cases, no apparent cause has been identified. This project aims to identify areas exposed to harmful RF radiation levels through the innovative application of blockchain technology. Our methodology involves measuring RF radiation in densely populated locations and implementing a blockchain-based spectral power misuse detection system via mobile networks. This system notifies users of harmful RF radiation exposure through an Android app. A clustering method and an aggregation method are included to simplify the blockchain process and make it more efficient. A consensus algorithm ensures the ledgers' consistency among all the blockchain peers. Our evaluation of the consensus algorithm focuses on performance metrics, including computational complexity, latency, and throughput. Additionally, smart contracts have been created to generate RF radiation patterns specific to particular areas, which telecommunication regulators can utilize to manage and regulate RF radiation effectively.

Contents

List of Figures

List of Tables

Acronyms

RF	-	Radio Frequency
SDR	-	Software Define Radio
RSL	-	Received Signal Level
SNR	-	Signal to Noise Ratio
PoS	-	Proof of Stake
2G	-	2nd Generation
3G	-	3rd Generation
4G	-	4th Generation
5G	-	5th Generation
TRCSL	-	Telecommunication Regulatory Commission in Sri Lanka
ICNIRP	-	International Commission on Non-Ionized Radiation Protection
WHO	-	World Health Organization
IARC	-	International Agency for Research on Cancers
GPS	-	Global Positioning System
FDA	-	Food and Drug Administration
DNA	-	Deoxyribonucleic acid
NTP	-	National Toxicology program
DBSCAN	-	Density-Based Spatial Clustering of Application with Noise
NDM	-	Network Data Mining
UI	-	User Interface
FCM	-	Firebase Cloud Messaging
AES	-	Advanced Encryption Standard
DES	-	Data Encryption Standard
SDK	-	Software Development Kits
WCSS	-	Within Center Sum of Square
MAD	-	Median Absolute Deviation
IQR	-	Interquartile Range
RANSAC	-	Random Sample Consensus
IQR	-	Interquartile Range
MAD	-	Median Absolute Deviation

DApps	-	Decentralized Applications
DLT	-	Distributed Ledger Technology
CFT	-	Crash Fault Tolerance
BFT	-	Byzantine Fault Tolerance
FCM	-	Firebase Cloud Messaging
SDKs	-	Software Development Kits
TPS	-	Transactions Per Second
MB	-	Megabytes

Chapter 1

Introduction

With the rapid increase in technology, the demand for all telecommunication services has increased considerably within the past few decades. With that increase in demand, telecommunication operators are attempting to improve the coverage of their 2G (2nd Generation), 3G (3rd Generation), 4G (4th Generation), and 5G (5th Generation) networks by adding more base stations, which generally are shared by multiple operators. These base stations are often located in urban areas. Therefore, it is essential to measure the level of radio frequency exposure for people close to these stations, typically within a few hundred meters, to ensure that it is safe for human exposure.

1.1 Background

The Telecommunications Regulatory Commission of Sri Lanka (TRCSL) is the national regulatory agency for the telecommunications industry in Sri Lanka. It follows the standards defined by the International Commission on Non-Ionized Radiation Protection (ICNIRP) [1] to issue the regulations for local service providers. However, some countries have taken additional measures beyond these standards to limit Radio Frequency (RF) exposure. Additionally, ongoing research is to determine whether humans may experience adverse health effects due to higher levels of RF radiation from mobile towers [2].

This research work aims to study the levels of RF radiation to which the general public is exposed in different regions of the country, focusing on locations where children and the general public frequently spend time. The measured RF radiation levels will be compared to accepted exposure limits in Sri Lanka. The study also aims to use blockchain technology integrated into the mobile network to detect misuse of RF spectrum power and to develop a mobile application to alert users about high levels of RF radiation in their living areas. This research could potentially be used to revise the current radiation exposure limit standards and tower policies in Sri Lanka.

Blockchain is a decentralized database technology used as a distributed ledger. It allows multiple parties to securely record and verify transactions without relying on a central authority. The database is transparent, secure, resilient, decentralized, and maintained by a network of interested parties who want to keep the ledger up to date. Each node in the network can maintain a copy of the ledger and control its data. Everyone sees the same ledger. There is no need for a master copy because the entries in the ledger are validated as they are entered using a process called consensus, which is implemented using various algorithms. Smart contracts can also be executed on the blockchain for time-sensitive activities. Therefore, approaches to spectrum management using blockchain technology have been proposed in the literature [3], [4].

1.2 Problem Statement

In the past few years, the rate of human deaths due to different cancers has increased exponentially in Sri Lanka [5]. The reasons for these are predicted as using instant foods, using low-quality plastic containers, and many more [6], [7]. It is also believed that exposure to higher RF levels in mobile towers harms humans, and scientists doubt that human exposure to higher RF levels might cause cancers. Also, RF radiation has been classified as a "possible human carcinogen" by the World Health Organization (WHO) [8] and the International Agency for Research on Cancer (IARC) [9]. Furthermore, uncontrolled radiation of RF affects pre-adolescent children, pregnant women, elderly humans, and patients with pacemakers. Therefore, detecting spectrum power misuse and alerting the public and telecommunication regulators about higher RF radiation is essential.

1.3 Aim

The aim of this project is to develop an RF radiation exposure level measurement system that can be used to analyze the RF radiation in different areas and notify the users exposed to harmful signal levels.

1.4 Objectives and Scope

1.4.1 Objectives

The main objectives of this project are designing and implementing an RF radiation exposure level measurement system using a blockchain platform for data storage and

developing a mobile application for accessing and visualizing recorded data. The specific objectives under the main objective mentioned above can be listed below.

1. Conduct the RF radiation power measurement survey on densely populated places.
2. Develop a mobile application to alert users about the higher RF radiation in their areas.
3. Implement the blockchain-based spectral power violation detection mechanism for mobile networks.
4. Provide a dynamic map showing the areas exposed to higher RF radiation levels to the mobile users and the telecommunication regulator.

1.4.2 Scope

1. Conducting the spectrum power measurement survey on five identified urban, suburban, and rural locations. These areas include schools, hospitals, and markets in urban, suburban, and rural areas.
2. The mobile application is only implemented on the Android platform. The application records the radio frequency power level in the area and transmits it to the base stations and miner nodes in the blockchain.
3. Integrate blockchain technology into the system to store and manage the collected data. The nodes in the blockchain network are the base stations of telecommunication network providers, called edge devices.
4. Dynamic map highlights the areas exposed to higher RF radiation levels, and the exact RF level of that location can be observed.

Chapter 2

Literature Review

Due to its possible effects on human health, Radio Frequency radiation released by wireless communication devices is receiving more and more attention.

The Journal of Endocrinology Research carries a study on processed radio frequency and pancreatic health that explores its impact on human health, particularly its potential relationship with diabetes [10]. This study highlights the possibility that wireless sensor networks lead to frequent urination and infections. The risk of diabetes is particularly alarming for overweight individuals exposed to these networks in dimly lit environments. The study emphasizes the value of dynamic sensor networks in healthcare while highlighting some drawbacks, such as the possibility of Global Positioning System(GPS) based digital poisoning. [10].

When considering RF Radiation and Cancer Risk, another study discussed in this review addresses the historical evidence concerning the risks of RF radiation, especially regarding cancer prevention [2]. This paper emphasizes that the various studies completed over many years have shown a higher risk of cancer linked to exposure to RF radiation. Inadequate research on the health and environmental effects of the fifth-generation wireless communication technology also raises concerns regarding its implementation. This study underlines the significance of an unbiased scientific evaluation before deploying technologies emitting RF radiation. [2].

Another study includes an evaluation of the health effects of RF radiation as determined by the United State's Food and Drug Administration (FDA) during an analysis of RF radiation's health impact[11]. While acknowledging the heating effects of RF radiation at high exposure intensities, this paper reviews the hypothesized non-thermal effects on biological processes. Regardless of the lack of evidence for direct DNA (Deoxyribonucleic acid) damage or tumor induction at non-thermal levels, the FDA nominated cell phone RF radiation to the National Toxicology Program (NTP) for a comprehensive assessment of potential toxicity and carcinogenicity [11]. This paper emphasizes the importance of well-designed animal experiments to assess the long-term health effects of RF radiation.

Furthermore, in light of animal studies' relevance to human health risks, the review delves into the importance of animal models in assessing health risks from environmental agents [11]. It emphasizes biological similarities between humans and animals and that recognized human carcinogens have shown evidence of carcinogenicity in animals. The National Toxicology Program conducted studies on cell phone radiation, revealing significant findings, including tumor and genotoxicity data, reduced pup birth weights, and cardiomyopathy induction in rats. These findings invalidate that low-level cell phone radiation cannot harm health. [11].

Collectively, these studies underscore the potential health implications of RF radiation exposure. The effects of RF radiation on human health require thorough investigation due to the correlation between it and diabetes, the increased risk of cancer, and the requirement for reliable animal research. The need for comprehensive evaluation before deploying technologies emitting RF radiation is evident, particularly in the context of 5G implementation. As research continues, understanding RF radiation's effects on human health remains a critical pursuit.

Blockchain technology has attracted significant attention recently for its potential to revolutionize various industries through its decentralized nature, cryptographic security, and peer-to-peer transactions [3][4]. Blockchain was first introduced by Satoshi Nakamoto in 2008 as a method to enable decentralized peer-to-peer payment transfers. Still, it has since developed into a flexible technology with various uses [3]. This study includes the advantages of blockchain technology and its potential impact on multiple sectors, including spectrum management.

When analyzing the benefits of blockchain technology, its ability to provide transparency and security in transactions is what motivates adoption [3]. Employing cryptographic algorithms, blockchain secures and verifies transactions, rendering data alteration or tampering virtually impossible. As a result, there is less chance of fraud or manipulation, and data integrity is ensured [3]. Additionally, the decentralized nature of blockchain eliminates the need for intermediaries, consequently reducing costs and enhancing efficiency across industries [3].

Another study proposes an efficient and effective blockchain-based data aggregation system for voting [12]. The authors mention that aggregation allows for collecting and consolidating data from multiple sources and databases, providing a comprehensive view of the overall count of candidates who participated in the election [12]. Therefore, it helps ensure the voting process's accuracy and integrity. Furthermore, the authors mention that aggregation also enhances data processing efficiency by distributing the workload among multiple peers, resulting in increased data processing frequency. The system utilizes blockchain technology to handle the vote and uses the secure hash algorithm to resolve security issues. Additionally, the authors discuss the importance of elections in a democratic society and the need for open and depend-

able voting systems. Overall, aggregation in the proposed system improves the voting process's security, trustworthiness, and privacy.

Smart contracts are a notable application of blockchain, which automates processes by executing predefined conditions [3]. These self-executing contracts eliminate intermediaries and mitigate the risk of human error or fraud, streamlining processes such as financial transactions, real estate agreements, and legal contracts [3].

When considering the clustering algorithms, this study focuses on clustering algorithms in the context of big data analytics and parallel computing [13]. It introduces the Fast Density-Grid Clustering Algorithm, designed to handle large datasets with fast execution times efficiently. This algorithm partitions the data space into a grid and assigns density values to each grid cell. Clusters are formed by merging neighboring spaces with the highest density. Experimental results for the serial version show that it offers accuracy similar to DBSCAN (Density-Based Spatial Clustering of Applications with Noise) but with faster execution times.

Another study addresses the importance of mining patterns from event logs and presents a novel clustering algorithm for log file data sets [14]. The algorithm considers the specific properties of log file data, such as the occurrence of words and correlations between frequently occurring words. Furthermore, it uses a density-based approach to identify frequent words and build cluster candidates, which are then selected as clusters. The algorithm is superior to an Apriori-based clustering scheme regarding runtime cost. However, it may consume a lot of memory, which is a drawback of the method. Therefore, the authors propose a technique to estimate and exclude irrelevant words from the vocabulary [14].

The diverse applications of blockchain are evident across multiple sectors, including finance, healthcare, government, manufacturing, and distribution [4]. As blockchain permeates domains such as supply chain management, digital media transfer, remote service delivery, and distributed resources, its potential becomes increasingly apparent [4].

Another study explores the potential impact of blockchain technology on the telecommunications industry [15]. This study examines various blockchain types (public, permissioned, and consortium) and consensus methods (such as proof of work and proof of stake.) in telecom networks. It suggests multiple blockchain applications in telecom, such as identity management and IoT security, and discusses opportunities for smaller operators. Emphasizing the need for flexible regulations, it highlights blockchain's potential to reduce operational costs and enhance security and efficiency. The study underscores blockchain's transformative impact on telecommunications, offering innovative solutions and market disruption possibilities.

When considering the challenges and considerations, Notwithstanding its potential, blockchain technology faces particular challenges that need to be addressed [4].

Scalability, for instance, poses a constraint as evidenced by the limited block size and frequency in networks like Bitcoin [4]. Technical solutions are required to overcome these challenges and enhance blockchain's practicality in high-throughput scenarios [4].

Additionally, security vulnerabilities, including the potential of a 51 percent attack, and regulatory considerations necessitate careful attention [4]. These issues demand comprehensive solutions to ensure blockchain technology's sustainable adoption and implementation.

Therefore, we can conclude that Blockchain technology offers many advantages, including transparency, security, efficiency, and automation [3]. Its decentralized nature and cryptographic foundations provide tamper-proof solutions across various sectors [3], [4]. However, addressing challenges such as scalability and security remains essential for its widespread adoption [4]. The continued evolution of blockchain technology promises transformative potential across diverse domains, sparking ongoing exploration and innovation.

This study discusses the NDM (Network Data Mining) strategy for mitigating data redundancy in IoT-driven wireless sensor networks [16]. Network Data Mining involves data mining techniques to select suitable information for transmission to a base station or cluster head. It employs classifier rules and similarity scores to determine data packet forwarding. The study assesses NDM's performance via simulations and compares it with existing methods, focusing on energy consumption, delay, throughput, network lifetime, and redundancy. The results indicate that NDM surpasses other approaches, enhancing network efficiency, resource utilization, and the number of operational sensor nodes while reducing end-to-end delay. In summary, NDM offers an effective solution for data redundancy elimination, enhancing IoT-based wireless sensor network performance [16].

Chapter 3

Methodology

3.1 Research Methodology

1. Literature review on several areas such as various kinds of health issues happening due to higher RF radiation and basics and applications of Blockchain technology.

2. Conduct the RF radiation power measurement survey on densely populated places such as schools, hospitals, and markets in urban, suburban, and rural areas.

3. Developing a mobile application to measure and transmit the measured signal levels and location values of the mobile user to the server and sending an alert to the users in areas with harmful signal levels.

4. Modifying an aggregation mechanism to gather data from mobile devices and check the validity of the data. It includes a calculation to choose a suitable value that can represent the data set in each cluster.

5. Developing a blockchain network with suitable smart contracts to record and detect high RF radiation levels.

6. Implement the blockchain-based spectral power violation detection mechanism over mobile networks. This includes implementing suitable smart contracts.

3.2 RF Radiation Survey

Before conducting the radiation survey at the densely populated places as identified, we have to validate the measurements received from the mobile phones. For that, we conducted the lab test by comparing the measurement received by the mobile phone antenna and the omnidirectional antenna connected to a spectrum analyzer. The test setup is shown in Figure 3.1

Figure 3.1: Test setup for validating mobile phone reading.

From this validation test conducted by placing the mobile phone and omnidirectional antenna nearby, we validate that the mobile phone's received signal level reading is accurate enough for the application that gives alerts based on a high RF exposure level. This test compares the mobile phone reading and spectrum analyzer reading for two mobile service providers' frequencies, as shown in Figures 3.2 and 3.3. In the mobile phone received signal level reading, we have used a third-party app named "G-NetTracker Lite."

Figure 3.2: Comparison of the spectrum analyzer reading and mobile phone reading for Dialog download frequency 1850 MHz.

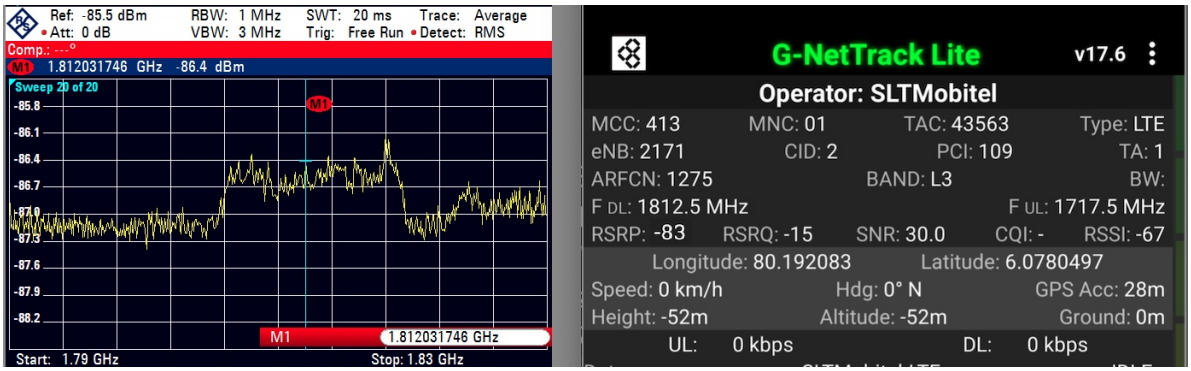


Figure 3.3: Comparison of the spectrum analyzer reading and mobile phone reading for Mobitel frequency 1812 MHz.

A survey was conducted to measure RF signal levels in different places to get an idea about the current RF signal emission from the base stations. Data was measured using the same mobile app used for the validation test. First, the survey was conducted from Hanthana Base station to Kandy town. Then, the second survey was done in the Galle district, considering areas like Galle Fort, around the Karapitiya Hospital, and around the Faculty of Engineering, University of Ruhuna. These data were analyzed using different aggregation methods. The Figures 3.4, 3.5, 3.6 and 3.7 show the data we collected from the survey at Kandy town, Galle Fort, Karapitiya, and Hapugala Engineering Faculty, respectively.

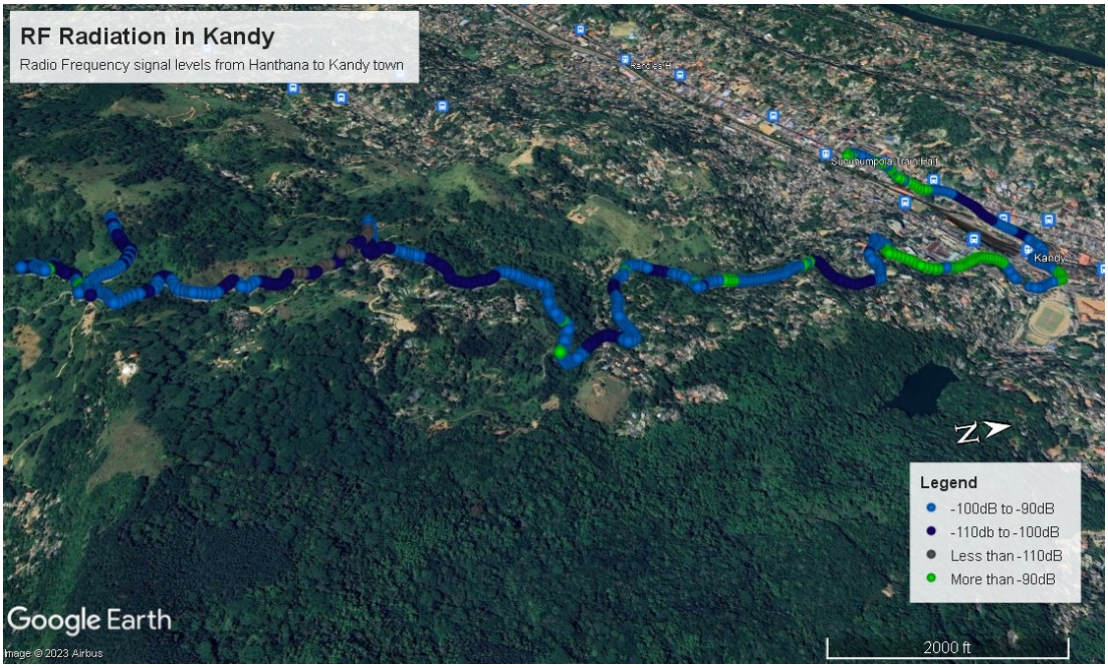


Figure 3.4: Representation of Data Obtained in Kandy District.

Figure 3.5: Representation of Data Obtained in Galle Fort.

Figure 3.6: Representation of Data Obtained in Karapitiya.

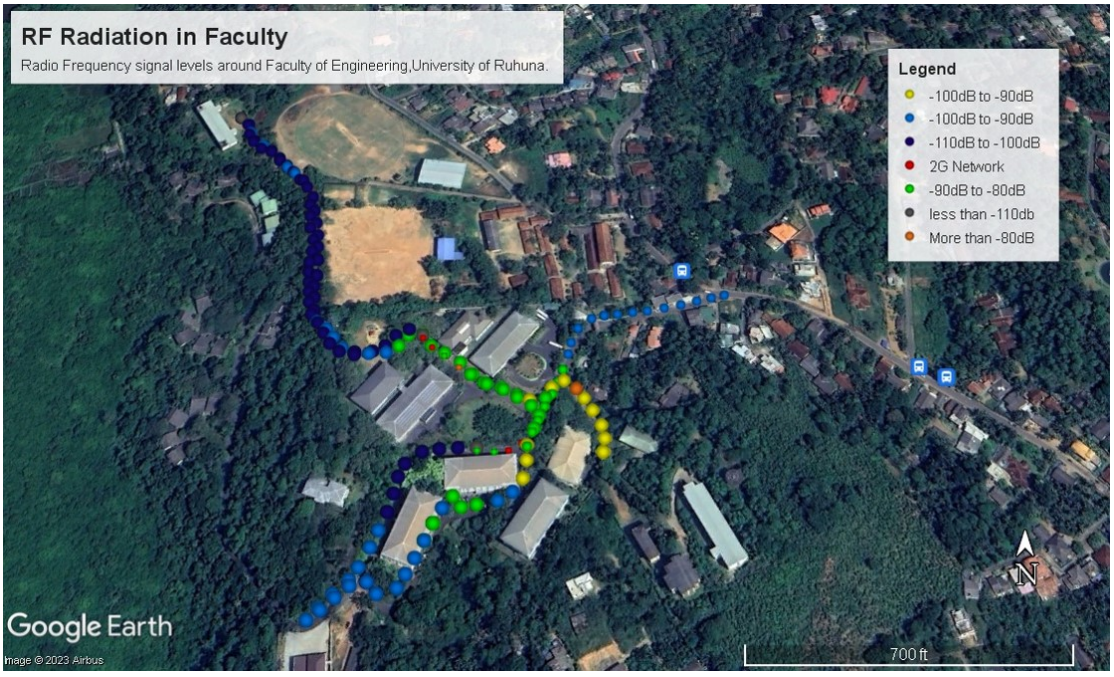


Figure 3.7: Representation of Data Obtained in Faculty of Engineering, University of Ruhuna.

3.3 Study on Public Awareness on Health Impacts of RF Radiation

This survey was conducted to gauge people's awareness of the health impacts of higher Radio Frequency radiation and assess the demand for the mobile application we developed to notify users about higher radio frequency radiation levels in their current location. There were 235 responses in total.

The diagram below illustrates the respondents' awareness, as indicated by the survey, regarding the potential health risks associated with exposure to higher radio frequencies, including the risks of cancer, reproductive issues, and neurological problems [17]. 88.5% of the respondents knew about the health impacts of higher Radio Frequency radiation.

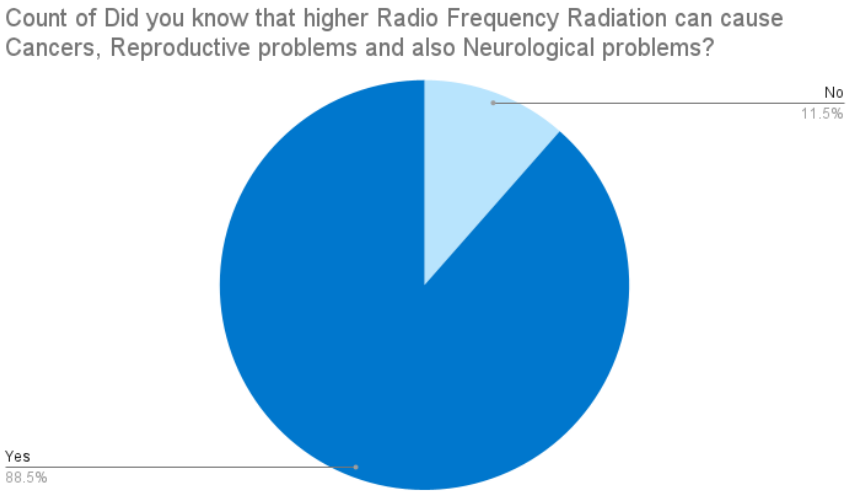


Figure 3.8: Public Awareness on Health Impacts of RF Radiation.

The diagram below illustrates the demand for mobile applications, that inform users about their exposure to higher Radio Frequency radiation. Impressively, 89.8% of the respondents expressed a favorable disposition toward the mobile application.

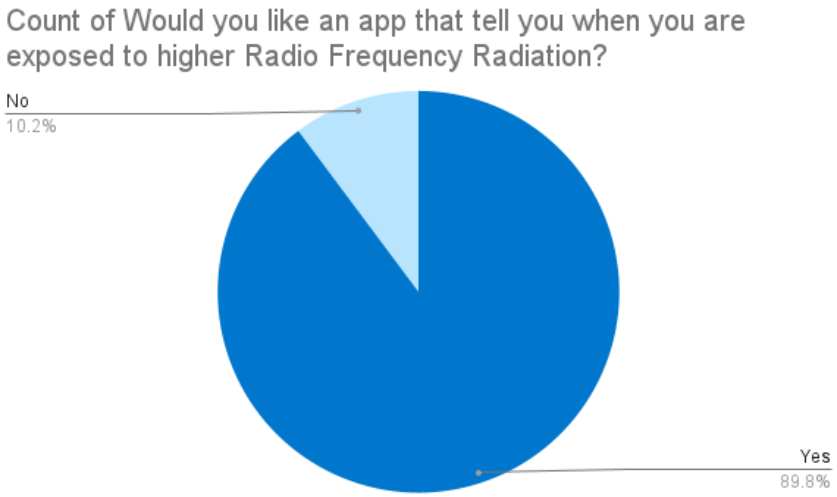


Figure 3.9: Demand for Mobile Application.

Based on the findings of the above survey, we developed a mobile application to detect and alert users about harmful RF signals.

3.4 Mobile Application Development

When developing the mobile application, instead of using the native development platforms for different operating systems, using mobile app development libraries has been a popular method lately. Flutter is one such popular open-source UI toolkit that allows developers to build high-quality native mobile apps for both Android and iOS platforms. Dart is the programming language that is used in Flutter. It's a modern, object-oriented programming language developed by Google. In our project, we developed an Android mobile application. We have chosen Flutter and Dart to create the mobile application, considering the following advantages.

1. **Fast Development**: Flutter's hot reload feature allows us to change our app's code and see the results immediately. Therefore, it makes the development process faster.

2. **Cross-Platform Development**: Flutter allows developers to build applications for both Android and iOS platforms using a single codebase. Therefore, it saves time and resources. In the future, if we want to expand the project further, we can easily build mobile applications for iOS platforms.

3. **Attractive and Customizable UI**: Flutter is a rich set of pre-built widgets and customizable UI components that help developers create attractive and engaging user interfaces.

4. **High Performance**: Flutter's high-performance rendering engine allows for fast and smooth animations. Therefore, it makes the user experience more enjoyable.

5. **Strong Community**: Flutter has a large and active community of developers and contributors. Therefore, it's easy to find help and resources when needed.

Because of these advantages Flutter provides, the Flutter UI framework alongside Dart programming language was selected to develop the mobile application of the project.

3.4.1 User Interface of the Mobile Application

Using the UI APIs provided by Flutter, a simple, intuitive, and user-friendly user interface was developed for the mobile application. In this, the current user location, the signal level, and a map with harmful radiation levels are displayed to the user. Figure 3.10 shows a screenshot of the mobile application showing the instantaneous signal level, the instantaneous location of the mobile user, and the map, which indicates the high signal levels in the area.

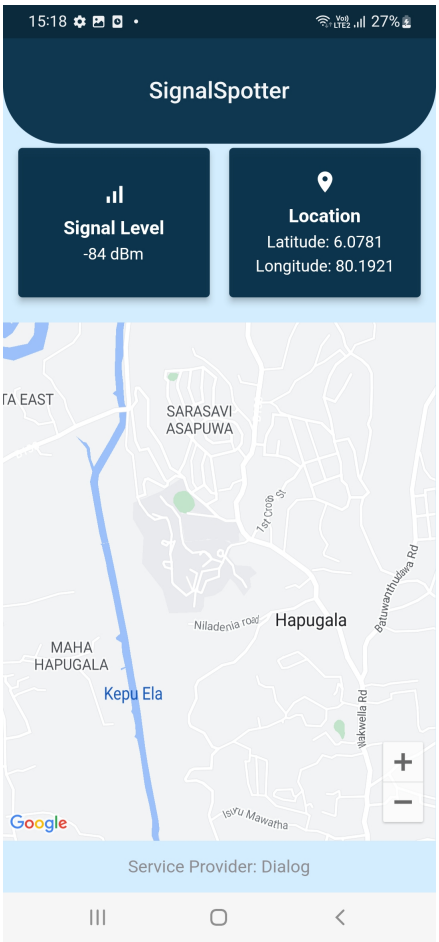


Figure 3.10: User Interface of the Mobile Application.

3.4.2 Data Transmission and Notification

The measured values from the mobile application are sent periodically to a server located in base stations. Figure 3.11 shows the log of data transmitted to the server.

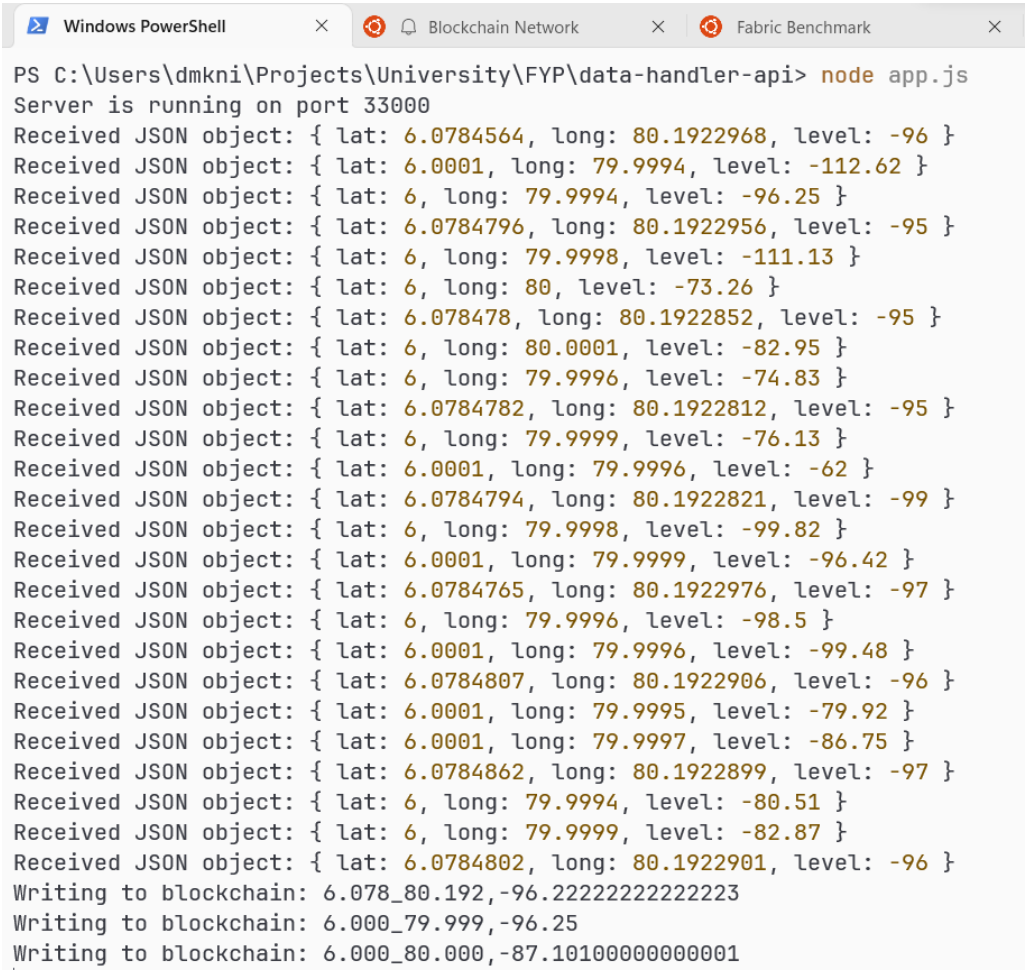


Figure 3.11: Data Transmitted to the Mobile Edge Device.

This data is then added to the blockchain, and from there, the data is analyzed, and a notification is sent to the users if a high RF radiation level is detected in their area. Figure 3.12 displays the notification sent to the user.

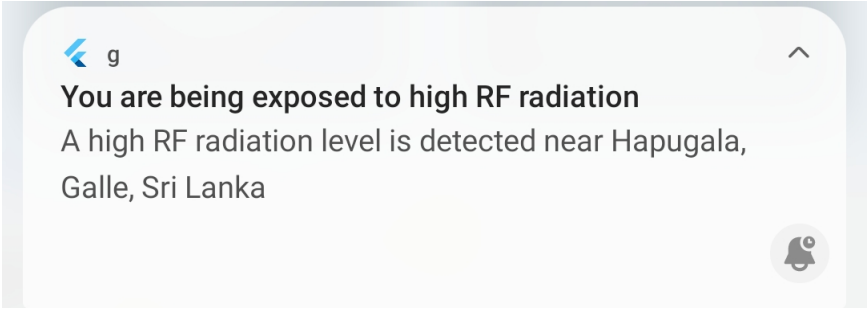


Figure 3.12: Notification after Detecting a Higher RF Level.

The identified hazardous RF emissions levels will be visually represented on a dynamic map within the mobile application. This feature will enable all mobile application users to access information regarding the geographic location with high RF radiation levels. This is graphically shown in 3.13, showcasing the dynamic map's functionality.

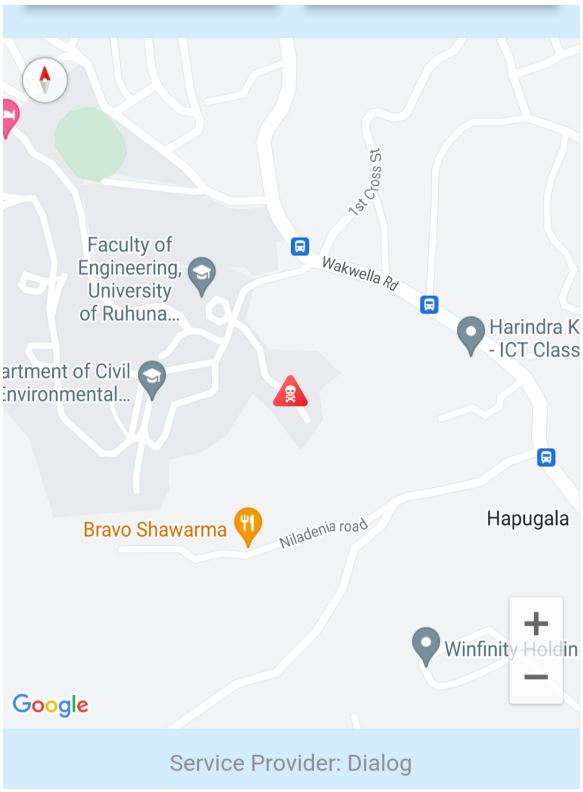


Figure 3.13: The Map Which Indicates High Radiation Level.

Securing Data Transmission

Since the mobile app transmits the mobile users' location data, an additional security measure was taken to increase the users' confidentiality. To implement this, the measured data must be encrypted using a suitable encryption algorithm before being transmitted to the server. Since additional user data is not transmitted from the mobile device, using a resource-heavy encryption algorithm is unnecessary. Therefore, a symmetric key encryption algorithm can be used for this application since they do not require additional resources for tasks like public/private key generation and management of digital signatures, compared to the asymmetric key encryption algorithms.

When considering different symmetric key algorithms, the most famous algorithms are Advanced Encryption Standard (AES)[18] and Data Encryption Standard (DES)[19] algorithms. A comparison between these two popular encryption algorithms is given below in Table 3.1.

Table 3.1: Comparison of AES and DES Encryption Algorithms.

AES	DES
AES is a symmetric encryption algorithm adopted by the U.S. government as the standard for encrypting sensitive data. It is widely used for secure data transmission and storage due to its strong encryption capabilities.	DES is an older symmetric encryption algorithm that has been widely used. However, it is now considered obsolete and less secure due to its shorter key length.
AES supports key lengths of 128, 192, and 256 bits, providing a higher level of security with longer keys.	DES uses a fixed key length of 56 bits, which is considered short by today's standards and makes it vulnerable to brute-force attacks.
AES operates on data blocks of 128 bits, which enhances its security and resistance to various cryptographic attacks.	DES uses a block size of 64 bits, which is relatively small and can lead to vulnerabilities when encrypting large amounts of data.
AES is highly secure and resistant to known cryptographic attacks when used with appropriate key lengths. It is considered one of the most secure encryption algorithms available today.	DES, due to its short key length, is no longer considered secure against modern attacks, and it is not recommended for use in sensitive applications.

When observing the comparison of the two algorithms given in Table 3.1, we can see that the AES encryption algorithm offers more security than the DES encryption algorithm at the expense of some additional resource consumption. However, when comparing it with the asymmetric encryption algorithms, the AES encryption algorithm consumes fewer resources. Therefore, the AES encryption algorithm was identified as the most suitable encryption method for our application, and it was used to encrypt the data before transmitting them on public network channels. The data was again decrypted at the receiving end before being sent to the blockchain for further analysis.

3.5 Data Clustering and Aggregation

Clustering is a process of grouping the received data points. Data points closer to each other in physical space are assigned to the same cluster. In our project, a proper clustering mechanism is needed to group the RF data so that the system can easily detect the areas in which the harmful RF signals are transmitted and alert the users in that area about it.

Data Aggregation is a technique used to combine multiple data points into a single value. In our project, the mobile application will collect a large number of RF radiation levels that telecommunication providers in Sri Lanka provide. Our

project aims to identify whether these RF radiation levels are harmful or not to the community and to inform them if the RF radiation level is harmful. The data we collected from the mobile application may contain a large number of repetitions as well as data with a negligible difference. Testing every RF radiation level measured by the mobile application consumes more power, time, and resources. Furthermore, a broken mobile phone may send faulty data to the servers, which can be detected as harmful data from our system and provide fake details to the users. Therefore, our system use an appropriate aggregation method with proper outlier detection to overcome these challenges. Our target for using an aggregation method for the project is to eliminate the faulty data(outliers) sent by mobile phones to servers and combine multiple data points into a single value.

3.5.1 Selecting an Appropriate Clustering Method

There are different types of clustering methods. From those methods, we selected three commonly used clustering methods that can be applied to our system. Those three methods are,

1. **K-Means Clustering**

   This clustering method is commonly used in machine learning [20]. From this method, the data set can be partitioned into a predefined number of clusters, where each cluster contains data points that are more similar to each other than to points in other clusters. The K-mean clustering method can be explained below,

   (a) **Step 1**: Define the number of clusters(k) randomly select points equal to the given number of clusters and consider them centroids of the data set.

   (b) **Step 2**: Assign each data point to the nearest cluster with randomly selected centroids based on the distance of each data point to each centroid.

   (c) **Step 3**: Recalculate the centroids of each cluster by computing the mean of all data points assigned to that cluster and again repeat step 2.

   (d) **Step 4**: Repeat Steps 2 and 3 are iteratively repeated until the assignments of data points to clusters no longer change significantly or a maximum number of iterations is reached.

   At the end of this process, all the data points of the data set are assigned to k number of clusters, defined at the beginning. This method will become complex when the number of clusters is unknown. Before following the abovementioned steps, we must find the number of clusters. For that, the Elbow method can be used. In that method, calculate the WCSS (Within Center Sum of Square)

error for k=1,2,3,4,5... and then create the graph of WCSS error vs. the number of clusters. After that, select the number of clusters with the highest gradient difference and follow the steps above.

When considering the K-mean method, it is easy to understand and can handle many data sets efficiently. However, when considering the disadvantages of this method, it forces the data into a predetermined number of clusters, which might not accurately reflect the underlying structure of the data. It also assumes that clusters have roughly equal variance and are similar sizes. When considering our project, these disadvantages may affect the accuracy of the project. Also, it increases the complexity of analyzing RF signal levels, identifying harmful signals, and alerting users about harmful RF signals.

2. **Hierarchical Clustering**

Hierarchical clustering is a popular clustering method used in machine learning [21]. This method groups similar data points into clusters, forming a tree-like structure called a dendrogram. This method can be further divided into agglomerative (bottom-up) or divisive (top-down) methods. Let's consider the agglomerative hierarchical clustering method. The basic steps of the method are explained below.

- Step 1: Consider all the data points as individual clusters.
- Step 2: Measure the distance between every cluster combined with the two clusters with the shortest distance and merge them into a single cluster.
- Step3: Repeat step 2 until it reaches the constraints given by the user, such as the number of clusters needed or the maximum distance between clusters.

The divisive hierarchical clustering method is the opposite of the agglomerative hierarchical method. In the divisive clustering method, all the data points are assumed to be single clusters, and then that cluster is divided into sub-clusters until it reaches the constraints.

Hierarchical clustering doesn't require a specific number of clusters. This method is flexible, where we can customize the constraints according to our requirements and cluster the data points according to those requirements. However, considering the disadvantages, this method can become computationally expensive for larger data sets as it is suitable for smaller ones. However, our system has to handle a large amount of data, and this method will increase its complexity. Another disadvantage of this system is that once the clusters are merged or split, it is challenging to modify the hierarchy, which can affect our project's accuracy.

3. **Grid-Based Clustering**

   Grid-based clustering is a common technique used to partition data space into a grid or lattice of cells, where each cell represents a small region within the data space. This method can be easily implemented on large data sets efficiently and scalable. Grid-based clustering simplifies the complexity of clustering algorithms by working with a predefined grid structure. Also, this method reduces the system's complexity by reducing the number of data points that need to be directly compared and clustered. This method is easily understandable, and the computational complexity is low. This can significantly speed up the clustering process, especially for large data sets with many data points.

Comparing the three clustering methods explained above clearly proves that the best option for the clustering method that can be used in our project is the Grid-based clustering method. We chose the Grid-based clustering method for our project and customized it according to our requirements.

3.5.2 Selecting an Appropriate Aggregation Method

Five main aggregation methods can be used to aggregate data samples collected by the mobile application [22],[23]. They are,

1. Median Absolute Deviation (MAD)
2. Interquartile Range (IQR)
3. Z-score
4. Tukey's method
5. RANSAC (Random Sample Consensus)

MAD method, IQR method, and Z-score method can be used only for the normally distributed data set. However, the data we collected from the mobile phone is not normally distributed; hence we have to convert that data set into a normally distributed one. This will reduce the system's accuracy and cause false detection in the community. Therefore, these three methods will not be used in our system.

Tukey's method and the RANSAC can aggregate data in our system.

Tukey's Method

This method is based on a combination of the interquartile range and the median of the data set. It is more robust than the other methods as it does not assume normality in the data. This makes it a better option for data sets that are not

normally distributed. Secondly, Tukey's method is more sensitive to the presence of outliers and can effectively identify and remove them, resulting in more accurate results. Another advantage of Tukey's method is its ease of understanding and simple implementation. Furthermore, it is appropriate for balanced and unbalanced designs, making it suitable for data sets with different sample sizes. The data set we obtain from the mobile application can vary in size and will not be a normally distributed data set. Therefore, this method for aggregating the data set will be more accurate than the other methods.

Steps of Tukey's method and the results obtained by applying this method to a randomly generated data set are given below,

1. Calculate the 25th percentile (Q1), the 50th percentile (median), and the 75th percentile (Q3) of the dataset.

2. Calculate the IQR as the difference between Q3 and Q1.

3. Define the lower bound as Q1 - 1.5 * IQR and the upper bound as Q3 + 1.5 * IQR.

4. Any data point that falls outside the lower and upper bounds is considered an outlier

Figure 3.14 shows the results obtained by performing Tukey's method to a random data set.

```
Command Window

>> tukey
The 25th percentile is -77.03
The 50th percentile is -65.89
The 75th percentile is -39.49
The interquartile range is 37.54
Lower Bound: -133.3344
Upper Bound: 16.8088
Filtered data:
  Columns 1 through 10

  -77.0307  -83.4054  -28.8891  -83.9363  -64.2651  -26.8278  -39.4949  -58.2817  -87.0255  -87.5422

  Columns 11 through 20

  -31.0329  -87.9415  -87.0017  -53.9763  -76.0196  -29.9320  -49.5233  -84.1015  -75.4545  -87.1645

  Columns 21 through 30

  -65.9018  -22.4998  -79.4391  -85.3795  -67.5115  -73.0418  -72.0193  -47.4559  -65.8835  -31.9831

  Columns 31 through 40

  -69.4232  -22.2283  -39.3846  -23.2320  -26.7992  -77.6420  -68.6380  -42.1970  -86.5155  -22.4112

  Columns 41 through 50

  -50.7121  -46.7091  -73.5862  -75.6640  -33.0811  -54.2835  -51.1910  -65.2419  -69.6555  -72.8281

Data after removing values beyond the boundary:
  Columns 1 through 10

  -77.0307  -83.4054  -28.8891  -83.9363  -64.2651  -26.8278  -39.4949  -58.2817  -87.0255  -87.5422

  Columns 11 through 20

  -31.0329  -87.9415  -87.0017  -53.9763  -76.0196  -29.9320  -49.5233  -84.1015  -75.4545  -87.1645

  Columns 21 through 30

  -65.9018  -22.4998  -79.4391  -85.3795  -67.5115  -73.0418  -72.0193  -47.4559  -65.8835  -31.9831

  Columns 31 through 40

  -69.4232  -22.2283  -39.3846  -23.2320  -26.7992  -77.6420  -68.6380  -42.1970  -86.5155  -22.4112

  Columns 41 through 50

  -50.7121  -46.7091  -73.5862  -75.6640  -33.0811  -54.2835  -51.1910  -65.2419  -69.6555  -72.8281

Mode of the filtered data:
  -87.9415

Mean of the filtered data:
  -59.6277
```

Figure 3.14: Results Obtained From Tukey's Method.

RANSAC Method

This robust estimation algorithm can be used in RF radiation exposure level measurements to identify and exclude outliers in the data [24]. It can estimate the parameters of a model even when a significant portion of the data is corrupted. This makes it an excellent choice for real-world applications where data may contain considerable noise or outliers. The method identifies and removes outliers to produce a more accurate

model for most data. Another advantage of RANSAC is its versatility. It can be applied to various problems, including linear and non-linear regression. Furthermore, the RANSAC method is automated and does not require any manual intervention. It randomly selects data points and determines the best model that fits most data. This makes it a convenient and time-saving method for fitting models to data sets, especially when dealing with large data sets.

Steps of the RANSAC method and the results obtained by applying this method to a randomly generated data set are given below,

1. Define a model that describes the expected relationship between the input variables and the output (e.g., the exposure level).

2. Randomly selects a subset of the data and uses it to fit the model.

3. Evaluate the model's fit to the rest of the data by calculating the residuals (the difference between the observed and predicted values).

4. Identify the data points with large residuals as potential outliers and exclude them from the data set.

5. Repeat steps 2-4 until the best-fit model is found, as determined by some stopping criterion (e.g., the number of iterations).

Fig. 3.15 illustrates the outlier detection of the RANSAC method to a random data set. The data points with a value greater than the inlier tolerance (t) are detected as outliers, and those with a value lesser or equal to the inlier tolerance(t) are detected as inliers.

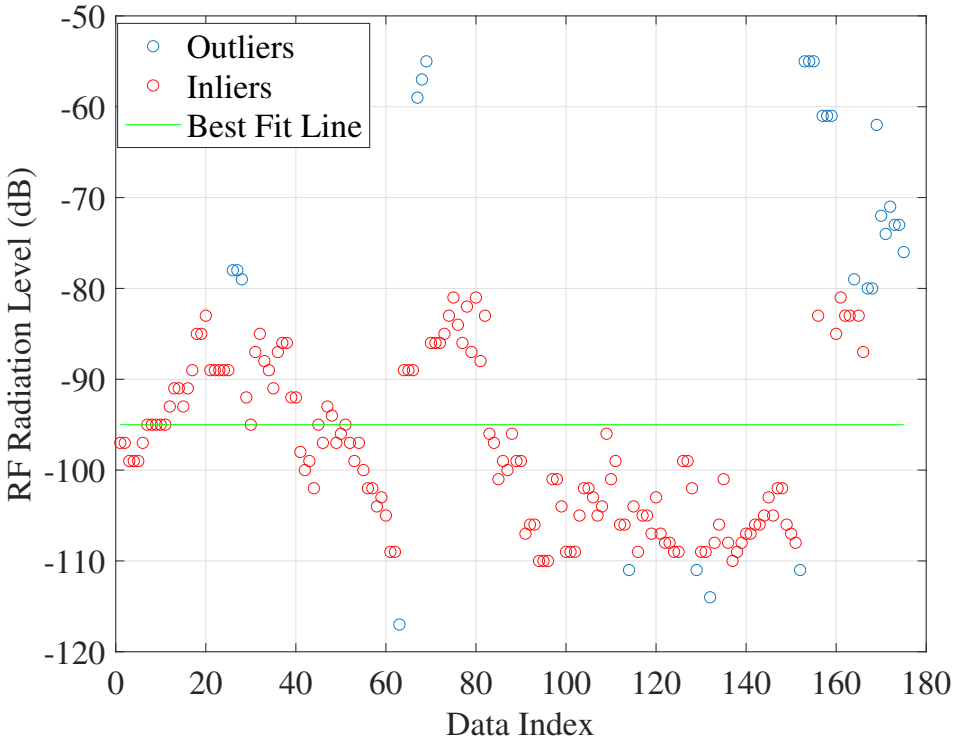


Figure 3.15: Outlier Detection Using RANSAC Method.

From these two methods, we chose the RANSAC method for our project.

Why We Choose RANSAC?

Out of the aggregation mechanisms discussed in Section 3.5.2, the RANSAC method was selected for our project to detect outliers and aggregate for the following reasons.

1. **Robustness to outliers**: RANSAC is designed to identify the outliers of a data set, making it suitable for our project to identify harmful RF signals that deviate significantly from normal behavior. But when considering Tukey's method it relies on medians for its calculations. While medians are less sensitive to outliers compared to means, they might not accurately capture the distribution of harmful RF data if the data has complex or asymmetric patterns.

2. **Adaptability:** The RANSAC method can be easily customized to fit different types of models, making it flexible for various scenarios. This method can customize parameters like the number of iterations, model order, inlier tolerance, and the number of points to create the model according to our project requirements. This adaptability allows one to choose the model that best represents the behavior of average RF data. However, Turkey's method is complex to implement and customize according to our requirements because it requires an understanding of statistical concepts and data analysis techniques.

3. **Computational Complexity:** When dealing with a large data set, RANSAC might be more suitable due to its potential for computational efficiency. In the RANSAC method, instead of using the entire data set for every iteration, a random sample of the data set is chosen, which paves the way for reducing the computational load significantly, as calculations are performed on a smaller subset of data. Additionally, it reduces the time and the memory required compared to processing the entire data set. However, Turkey's method might become computationally expensive for massive data sets due to the need to calculate medians, which are computed using the entire data set. It requires processing the whole data set for each iteration, which consumes more memory and time.

3.6 Blockchain Development

A blockchain is an immutable transaction record in a peer-distributed network. Bitcoin pioneered this, followed by Ethereum, which introduced smart contracts for apps. These public blockchains are open and support anonymous interactions. However, our application, where the users' location data are stored, requires a more advanced system with high confidentiality. *Hyperledger Fabric*[25] is an open-source, enterprise-grade permissioned distributed ledger technology (DLT) platform. Therefore, in our project, the Hyperledger Fabric DLT platform stores and analyzes the data in a distributed ledger.

Hyperledger Fabric was purpose-built for businesses, focusing on participant identification, permissioned networks, and transaction privacy. Some of the key reasons we selected Hyperledger Fabric over other popular blockchain platforms are listed below.

1. **Enterprise-Grade Technology**: Hyperledger Fabric is designed as an open-source, enterprise-grade distributed ledger technology platform. It is tailored for use in enterprise contexts, offering features that cater to the needs of businesses.

2. **Linux Foundation Backing**: Unlike many other blockchain platforms, Hyperledger was established under the Linux Foundation. This foundation is renowned for nurturing successful open-source projects, fostering strong communities, and ensuring open governance.

3. **Diverse Governance**: Hyperledger Fabric operates under a diverse technical steering committee and project maintainers from various organizations. It boasts a rapidly growing development community of over 35 organizations and nearly 200 developers.

4. **Modular and Configurable**: Fabric boasts a highly modular and configurable architecture. This flexibility allows it to be adapted to a wide range of industry use cases, from banking to healthcare to supply chain management.

5. **General-Purpose Smart Contracts**: Unlike platforms that rely on domain-specific languages, Hyperledger Fabric supports smart contracts authored in general-purpose programming languages like Java, Go, and Node.js. This reduces the learning curve for developers.

6. **Permissioned Network**: Fabric operates in a permissioned environment where participants are known and identified, offering higher trust than anonymous public networks.

7. **Pluggable Consensus Protocols**: Fabric supports pluggable consensus protocols, enabling customization to fit specific use cases and trust models. This adaptability enhances efficiency and performance.

8. **Customizable Privacy and Confidentiality**: Fabric's channel architecture and private data feature enable confidential transactions and data sharing within authorized subsets of participants, catering to business privacy needs.

9. **Decoupled Consensus**: The consensus component is logically decoupled from transaction execution and ledger maintenance. This flexibility allows for the use of well-established crash fault-tolerant (CFT) or byzantine fault-tolerant (BFT) consensus mechanisms.

10. **Performance and Scalability**: Fabric addresses performance concerns by executing transactions before finalizing their order. This execute-order-validate architecture allows parallel execution, enhancing throughput and scale.

3.6.1 Hyperledger Fabric Blockchain Network

Because of the modular nature of the Hyperledger Fabric blockchain network, Hyperledger Fabric offers several deployment methods, each with its own set of advantages and considerations. These methods include:

1. **Docker-Compose Network:** Docker Compose is a widely used method for deploying a Hyperledger Fabric network. It allows for easy setup and management of a multi-container environment where each container represents a different component of the blockchain network.

2. **Multiple Physical Hosts:** Deploying Hyperledger Fabric on multiple physical or virtual hosts can enhance security and resilience. This method often involves setting up individual nodes for each network component.

3. **Cloud Platforms:** Cloud providers like AWS, Azure, and GCP offer managed services for deploying Hyperledger Fabric networks, simplifying infrastructure management

When considering these most popular methods to develop a Hyperledger Fabric Network, deploying the network on a cloud platform was not feasible for the development stage of our project because of the cost limitations. Also, deploying the network on multiple physical hosts causes high hardware resource usage and difficulties in development. Therefore, for the project's development stage, the most cost-effective and feasible solution for deploying the network was to deploy it on a Docker-Compose network. Docker Compose allowed us to efficiently utilize available computing resources by running multiple containers on a single host, making optimal use of the hardware. This efficiency was crucial in ensuring that our network ran smoothly within the confines of our resource constraints. Also, this allowed us to simplify the development process and reduce the cost while still having the opportunity to move to a much broader deployment option later in the production stage.

In addition to the deployment method, we considered several other critical factors when implementing our Hyperledger Fabric network. These included defining the network's topology to align with our use case, selecting an appropriate consensus mechanism, addressing data privacy concerns through channels, and more. These details about the architecture of the network will be discussed further in Section 4.5.

Chapter 4

RF Radiation Exposure Level Mesurement System

4.1 Design and Implementation of the System

The overall architecture of the implemented high RF radiation level detection system is shown in Fig. 4.1.

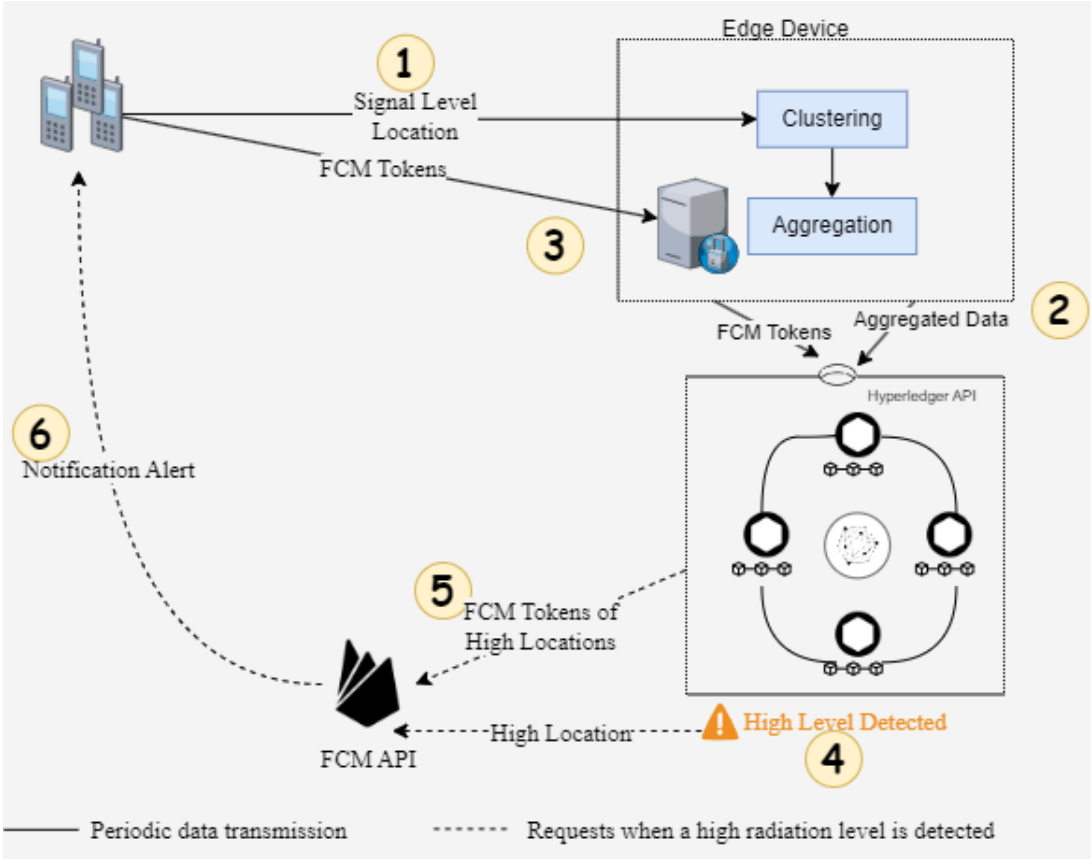


Figure 4.1: Overall Architecture of the System.

Here, the steps of operation of the system shown in Fig. 4.1 are as follows:

1. **Data sent from mobile devices:** The location data and the signal levels in that area are measured from the mobile devices and sent periodically to the base stations, from where they are clustered and aggregated.

2. **Data added to the Blockchain:** After clustering and aggregation, the data is added to the Hyperledger Fabric network for further analysis

3. **Firebase Cloud Messaging (FCM) Tokens are Added to the Blockchain:** The unique token of each mobile device is also updated periodically on the blockchain

4. **High RF radiation level is detected:** When a high RF radiation level is detected, a request is sent to Firebase Cloud Messaging API along with the location to send a notification to the users

5. **Relevant user tokens are queried:** When the request for a notification is received, the FCM tokens of users related to the given location are queried from the blockchain network

6. **Notifications sent:** After receiving the user tokens, the notifications are sent to the mobile devices with these tokens, alerting them that they are potentially being exposed to high RF radiation levels

Figure 4.1 gives a simple overview of the system's operation, and an in-depth explanation of each of these components is provided in the following subsections.

4.2 Developing a Mobile Application

As discussed in Section 3.4, the mobile application was developed using Flutter. The preceding discussion further explained and justified the reasons for choosing the Flutter platform. The overall architecture of the mobile application and other technologies used are described in this section.

Expected Features of the Mobile Application

1. Measure the signal levels in the area and display them.

2. Detect the current location in terms of longitude and latitude and display it.

3. Transmit the signal levels and the location to the server (Base Station).

4. Generate a dynamic map indicating the areas exposed to harmful RF levels.

5. Display an alert if the signal exceeds the threshold level. (notification includes about the higher and unhealthy signal level in the area in which the user is currently located)

Use of Firebase Cloud Messaging

Firebase Cloud Messaging was used with Flutter to implement the application's notification system.

Firebase is a comprehensive platform provided by Google that offers various tools and services for mobile and web application development. It covers multiple aspects such as authentication, real-time database, cloud storage, hosting, and more. Furthermore, Firebase simplifies the development process by providing developers with pre-built solutions that can be integrated seamlessly into their applications, reducing the need for extensive backend infrastructure setup.

Push notifications are essential for engaging users and informing them about important events, updates, or personalized content in mobile applications. In our project, we alert the users about the area's higher and more harmful radio frequency radiation levels. Firebase Cloud Messaging is a vital component of the Firebase platform that enables push notifications to be sent to mobile devices. Firebase Cloud Messaging is a reliable and efficient messaging solution that allows developers to easily reach users across different platforms, including Android and iOS.

Advantages of Using Firebase for Push Notifications:

1. **Easy Integration**: Firebase provides Software Development Kits (SDKs) for various platforms, making integrating push notification functionality into applications straightforward.

2. **Cross-Platform Support**: FCM supports multiple platforms, ensuring that notifications can be sent to various devices, irrespective of their operating systems.

3. **Reliability and Scalability**: Firebase handles the complexities of message routing, delivery, and retries, ensuring reliable message delivery. It is also built to handle large-scale applications with ease.

4. **Targeted Notifications**: Developers can send notifications to specific user segments based on user behavior, location, or preferences, enhancing user engagement. In our project, we send notifications to mobile phones in a particular area.

The use of the Firebase Notification system in our project is illustrated in Figure 4.2 below.

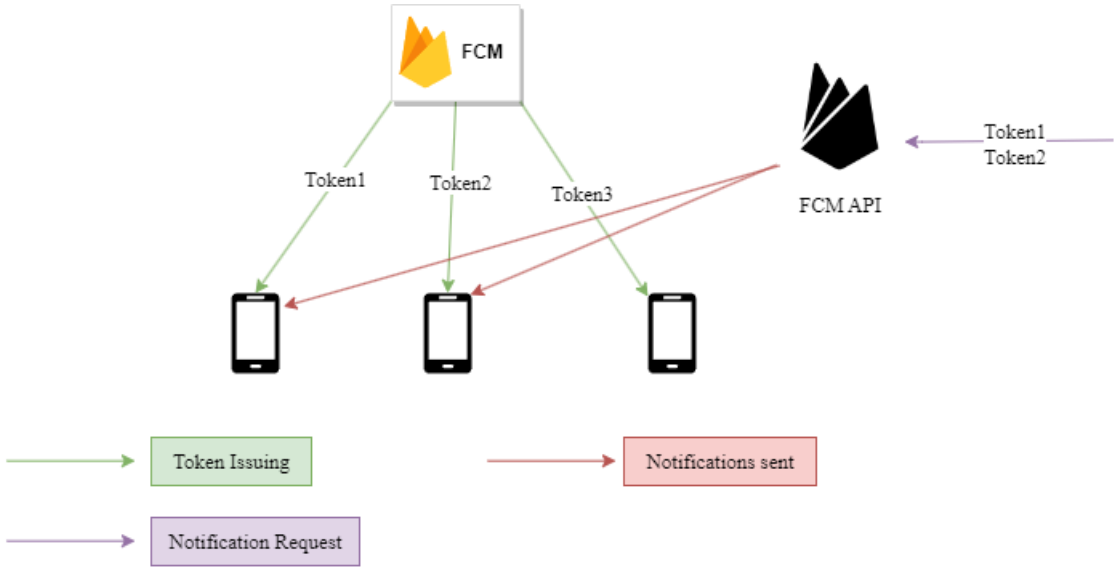


Figure 4.2: Firebase Cloud Messaging Architecture.

As Figure 4.2 shows, when the application is installed on a smartphone, a unique token is issued by the FCM for each device. Then, when a notification needs to be sent, FCM APIs can be configured to send required notifications to the devices identified by their unique tokens. This is the procedure used by the notification system in our application.

4.3 Grid-based Clustering Method

By considering the requirements of our project, we customized the Grid-Base clustering method and applied it to our project. Usually, the location of a particular place is given by using longitudes and latitudes. These longitude and latitudes are displayed using six decimal places—for example, latitude=2.324560, longitude=3.236786.A square-shaped area can be obtained when these values are rounded to some decimal places. This area varies with the number of decimal places which the location round off. The approximate cluster sizes we get from rounding latitude and longitude values are shown in Table 4.1.

Number of Decimal Points	Cluster Size
5	$1.1 \times 1.1\ m^2$
4	$11.1 \times 11.1\ m^2$
3	$111 \times 111\ m^2$
2	$1111 \times 1111\ m^2$

Table 4.1: Rounding Off Decimals and Corresponding Cluster Sizes.

Considering values shown in Table 4.1, it is demonstrated that when the number of decimal places decreases, the area of the obtained square increases. When consid-

ering our project, rounding off the location to 5 or 4 decimal places creates a small square, and using such a small square for our project wastes resources and increases the system's complexity. Furthermore, dividing the area of Sri Lanka into these small squares will create many squares, complicating the data analysis process and identifying the harmful RF signals. When considering the squared areas obtained by rounding off the location to 2 or 1 decimal place, It will create large squares, reducing the project's accuracy. As in the aggregation method, after removing the outliers, the average of the remaining inliers is taken as the value representing a particular cluster. However, when the data obtained in a specific cluster is large, it is difficult to take the average and detect the harmful signal levels. Therefore, considering all these scenarios, we choose the square with a $111 \times 111\ m^2$ area, which can be obtained by rounding off the longitude and latitude values to three decimal places for our project.

The below Fig. 4.3 shows the diagram of the area of the Galle fort, which is clustered according to $111 \times 111\ m^2$ squares.

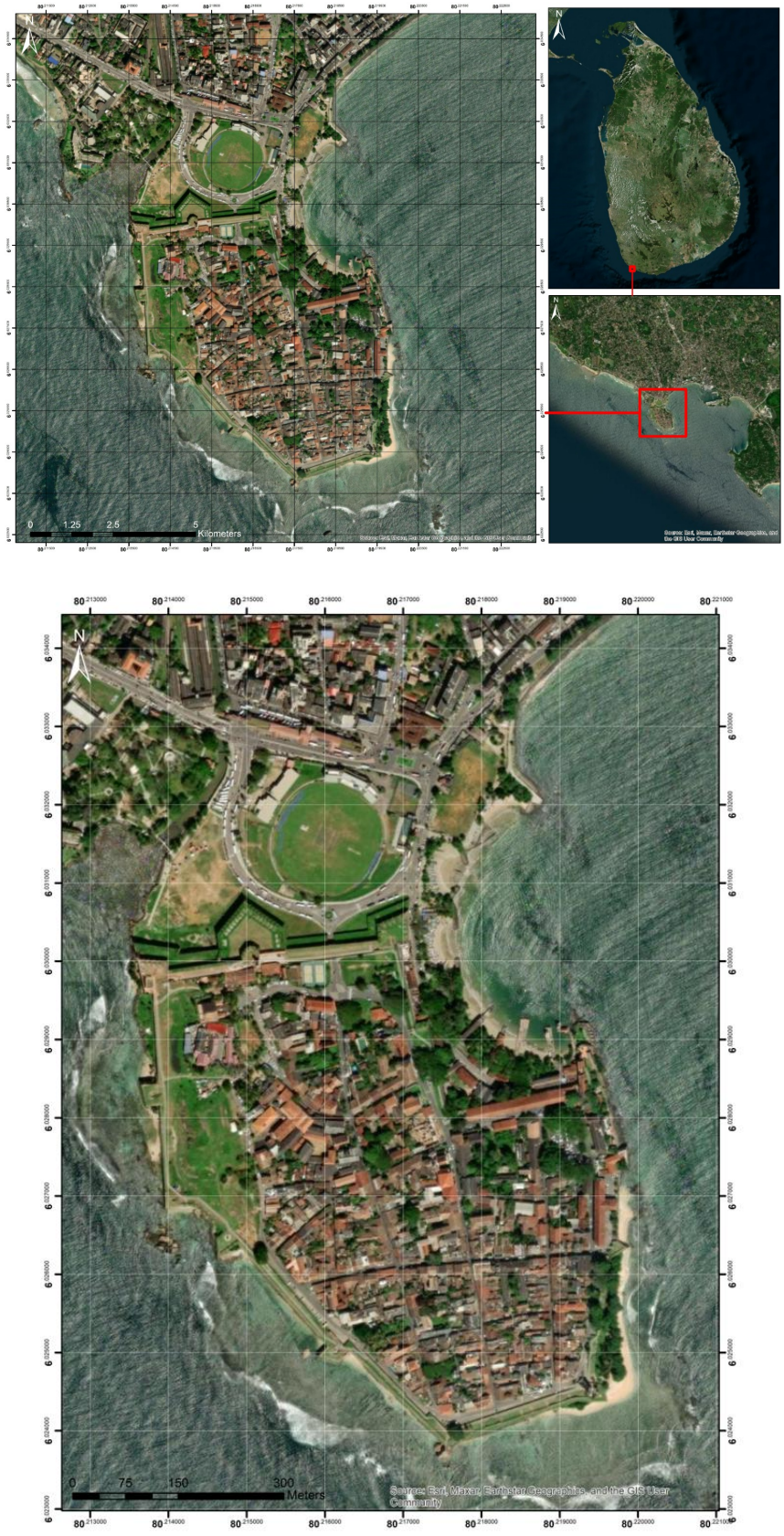


Figure 4.3: Grid Based Clustering Method for Cluster Size of 111×111 m^2.

4.4 RANSAC Method for Outlier Detection

Four main parameters should be entered for the RANSAC method. These parameters vary according to the project's different requirements. The four parameters are the number of points required to fit a line, the maximum number of iterations, Inlier tolerance, and the order of the RANSAC model we needed.

1. As the value for the number of points required to fit a line (n), two is chosen since it increases the number of combinations of data points to iterate through.

2. As the value for the maximum number of iterations (k),1000 is chosen because when the number of iterations increases. Furthermore, the accuracy of the system increases proportionally.

3. As the value of the inlier tolerance (t), 15dB is chosen for the simplicity of the calculation. This can be customized according to the preferences.

4. As the order of the RANSAC model, one is chosen because it gives a straight line by considering the computational simplicity.

The RANSAC method will be performed in our project based on the parameter values inserted based on the project's requirements. A brief explanation of the execution process of the RANSAC code, which is shown in Figure 4.4, is given below.

1. Select two random data points from the data set as the parameter of the number of points required to fit the line is given as two (n=2).

2. Model, a straight line of order one, using those randomly selected two points

3. Identify the inliers and the outliers based on the inlier tolerance value (t=15dB) and count the number of inliers.

4. Repeat 1,2 and 3 until the number of iterations reach 100(k=100)

5. Find the straight line with the highest number of inlier counts and select it as the best-fit model.

6. Remove all the outliers of the best-fit model and calculate the average of the inliers, and that will be the value that represents a data set of a particular cluster.

Fig. 4.4 shows the MATLAB implementation of the RANSAC method. This will detect the outliers and inliers as shown in Fig. 3.15. After removing outliers and calculating the mean of the remaining data points, we can obtain a value representing the whole data set in a cluster.

```
%RANSAC METHOD
% Data set
data = tdfread('data.txt','tab');

% RANSAC parameters
n = 2; % Number of points required to fit a line
k = 1000; % Maximum number of iterations
t = 15; % Inlier tolerance

best_model = zeros(1, 2);
best_count = 0;

for i = 1:k
    % Randomly select n points
    idx = randperm(length(data), n);
    sample = data(idx);

    % Fit a line to the sample
    model = polyfit(1:n, sample, 1);

    % Compute distance to the line for all points
    dist = abs(data - polyval(model, 1:length(data)));

    % Count inliers
    count = sum(dist < t);

    % Update best model if necessary
    if count > best_count
        best_model = model;
        best_count = count;
    end
end

% Identify inliers
dist = abs(data - polyval(best_model, 1:length(data)));
inliers = dist < t;
```

Figure 4.4: MATLAB Implementation of the RANSAC Method.

4.5 Blockchain Architecture

Because of the reasons discussed in Section 3.6, the Hyperledger Fabric Blockchain framework was chosen to implement the distributed network in our project. In this section, the architecture of the implemented decentralized network will be discussed in detail.

Architecture of the Hyperledger Fabric Network

Hyperledger Fabric is a highly modular and flexible blockchain platform for enterprise use cases. To effectively explain the network architecture we have used in this project, let's start by defining some key components:

1. **Organizations:** In the context of Hyperledger Fabric, organizations represent distinct entities that participate in the network. These organizations can be businesses, institutions, or entities that must transact on the Blockchain. For example, in our application's production environment, we can define organizations as separate telecommunication service providers.

2. **Peers:** Peers are nodes within the Hyperledger Fabric network that maintain copies of the ledger and smart contracts (chaincode).

   (a) Endorsing Peers: These peers simulate and endorse transactions by executing smart contracts. They are responsible for validating the proposed transactions before adding them to a block.

   (b) Committing Peers: Once transactions are endorsed, committing peers validate and commit these transactions into blocks. They maintain a copy of the ledger and are responsible for ensuring consensus is reached on the state of the Blockchain.

3. **Orderers (Ordering Service):** The ordering service, or simply "orderers," is responsible for grouping validated transactions into blocks and ensuring that these blocks are delivered to all the peers in the network in a consistent order. It's a crucial component for achieving consensus among network participants.

4. **Channel:** A channel in Hyperledger Fabric is a fundamental concept that enables private and confidential communication between a subset of network participants. It's like a private space within the blockchain network where only a specific group of organizations and peers can interact. Each Channel has its ledger, which records the history of transactions within that Channel.

5. **Chaincode:** Chaincodes, also known as smart contracts, are the executable programs that define the business logic and rules governing transactions on the Hyperledger Fabric network. They encapsulate the logic for validating, endorsing, and committing transactions to the ledger.

In our network, there are two separate organizations[1], representing two different service providers. Each organization operates independently, managing its own set

[1]This network configuration is defined for development purposes; they can be further improved in the production environment

of peers and resources. These organizations might have their own administrative policies and governance structures. For these organizations, we have included two peer nodes each, which can represent the base stations of the service provider. One Channel was created to connect all those four peers. Also, our network includes three orderers, which form the ordering service. The orderers work together to order and package transactions into blocks and then distribute these blocks to all peers in the network. Having multiple orderers enhances network fault tolerance and ensures high availability.

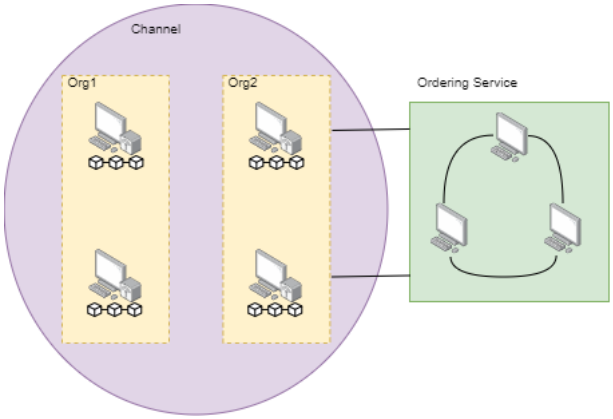


Figure 4.5: Architecture of the Fabric Network.

Figure 4.5 represents a simplified architecture of the network we have used in this project. Transactions within the network will follow a process where they are first endorsed by the endorsing peers, then ordered and committed with the consensus of the committing peers. Using channels ensures that not all network participants can access the same data, providing privacy and control over who can see and interact with the transactions on the Blockchain. This process will be further discussed under transaction flow.

Consensus Algorithm

In contrast to many distributed blockchains like Ethereum and Bitcoin, which are open to any node for consensus participation, Hyperledger Fabric adopts a distinct approach. These open blockchains depend on probabilistic consensus algorithms, which provide a high probability of ledger consistency but remain susceptible to ledger forks, where network participants have divergent views on transaction order.

Hyperledger Fabric, on the other hand, employs a specialized entity known as an 'orderer' (also referred to as an 'ordering node') to handle transaction sequencing. Multiple orderer nodes combine to form an ordering service. This design leverages deterministic consensus algorithms, ensuring that any block validated by a peer is definitively final and accurate. Consequently, the ledger cannot undergo forks,

a characteristic common in various other distributed and permissionless blockchain networks.

Hyperledger Fabric is known for its flexibility and modularity, and one of its key design principles is pluggable consensus. Fabric allows users to select and implement different consensus algorithms according to their requirements. The consensus algorithm defines how network participants agree on the order and validity of transactions. Fabric provides three main consensus options, which are,

1. **Solo**: The Solo implementation of the ordering service is intended for testing only and consists only of a single ordering node.

2. **Kafka**: Apache Kafka is a Crash Fault Tolerant (CFT) implementation that uses a "leader and follower" node configuration. Kafka utilizes a ZooKeeper ensemble for management purposes.

3. **Raft**: Raft is a CFT ordering service that follows a "leader and follower" model, where a leader node is elected (per Channel), and the followers replicate its decisions.

However, the Solo consensus algorithm is meant to be used with only one ordering node, and the additional administrative overhead of managing a Kafka cluster has caused the Kafka consensus algorithm to be intimidating and undesirable for many users. Therefore, both consensus algorithms have been deprecated from the latest versions of Hyperledger Fabric, and Raft has been the most popular consensus algorithm for production applications [26]. A further comparison between Raft and Kafka consensus algorithms is given in Table 4.2

Table 4.2: Comparison of Raft and Kafka Consensus Algorithms.

Raft	**Kafka**
Easier setup, embedded into ordering node	Complex setup, requires expertise in Kafka infrastructure and settings
Fewer components to manage	More components, more potential points of failure
Supports decentralization with separate ordering nodes for each organization	Centralized control in a single Kafka cluster
Native support with Fabric developer community backing	Requires users to obtain and manage Kafka and ZooKeeper images
Allows specifying ordering nodes for each Channel	Admin-controlled node allocation in Kafka
First step towards BFT ordering service	No direct support for BFT, separate transition process

Raft achieves consensus by electing a leader among a group of nodes (in your case, the orderers) responsible for ordering and packaging transactions into blocks. Here's a simplified explanation of how the Raft consensus algorithm works:

1. **Leader Election**: The Raft algorithm starts with a leader election process. Initially, all orderer nodes in your network are candidates for leadership. They send "heartbeat" messages to their peers to establish their authority. Once most nodes (more than half) acknowledge the leader's heartbeat, that node becomes the leader for a specified term.

2. **Log Replication**: The leader receives and orders incoming transactions into blocks. Once a block is formed, it is sent to all other orderer nodes for validation and replication. These orderer nodes replicate the leader's log of transactions to ensure redundancy and fault tolerance.

3. **Commitment**: To achieve consensus, a transaction must be included in the leader's log, replicated to most nodes, and then acknowledged back to the leader. Only when the leader receives acknowledgments from most nodes does it commit the transaction to the Blockchain, ensuring that it is part of the final ledger.

4. **Fault Tolerance**: Raft is designed to handle node failures gracefully. A new leader election is triggered if the leader fails or becomes unreachable, ensuring that the network continues operating despite node failures.

5. **Consistency**: Raft ensures that all nodes in the network eventually reach a consistent state. Once a transaction is committed by the leader and replicated to most nodes, it is considered final and immutable.

Chaincodes (Smart Contracts)

Chaincode, also known as smart contracts in Hyperledger Fabric, encapsulates the business logic of the blockchain application. In our Fabric network, two essential chaincodes, 'level' and 'token,' have been deployed within the common Channel shared among all peers. Together, they form a robust radiation monitoring and notification solution.

The 'level' chaincode serves the purpose of managing and analyzing the signal strength information. The signal level data measured from the mobile applications are clustered, aggregated, and transmitted to the Blockchain. These transmitted data are captured by the 'level' chaincode. This maintains comprehensive data about signal levels across various geographic locations. Also, the 'level' chaincode analyses the incoming data to determine if there are any locations with harmful radiation levels. Furthermore, a map indicating the locations with high RF radiation levels is

implemented on the mobile application. For this, the locations with high RF radiation levels are queried from the Blockchain using this 'level' chaincode.

The 'token' chaincode serves the purpose of managing the data of locations of mobile application users. This data is used when a high RF radiation level is detected in a particular area, and a push notification needs to be sent to the users in that location. The 'token' chaincode periodically records the users' locations, along with a unique token for each mobile device it is assigned to. Whenever a high RF radiation level is detected within the 'level' chaincode, the mobile application users in that particular location are queried using the 'token' chaincode, and the push notification is sent to all those users. An illustration of this notification process was shown and described in Figure 4.1.

Chapter 5

Results and Conclusions

5.1 Results

In this research, a Hyperledger Fabric blockchain network stores the RF radiation data and analyzes them for detecting high RF radiation levels. In this proposed architecture, the flow of operation consists of four main processes

1. *Data collection* using the mobile device
2. *Data clustering* based on location
3. *Data aggregation* to eleminate outliers
4. *Recording data* to the blockchain

The data clustering and aggregation steps are essential to maintain better network performance. By doing this, the outliers can be neglected, and the blockchain's throughput can be improved. Therefore, the performance analysis is done on the Hyperledger Fabric network by considering two cases: without clustering and aggregation mechanisms and with clustering and aggregation mechanisms.

We employed Hyperledger Caliper as the benchmarking tool for the performance analysis, assuming a user base of 1250 for the mobile application. Subsequently, 10,000 requests are dispatched to the blockchain network to evaluate performance metrics, including latency, throughput, and success rate. In both scenarios, transactions are transmitted over 2 minutes, each with a corresponding Transaction Per Second (TPS) rate, to observe the growth of the chain.

5.1.1 Case 1: Without Clustering and Aggregation Mechanisms

When no aggregation and clustering method is used before sending data to the blockchain, the number of TPS on the blockchain is significant since all the recorded

data points are sent to the blockchain. Considering the configurations of the mobile application, there are 250 TPS on the blockchain when the number of mobile phone users equals 1250.

5.1.2 Case 2: With Clustering and Aggregation Mechanisms

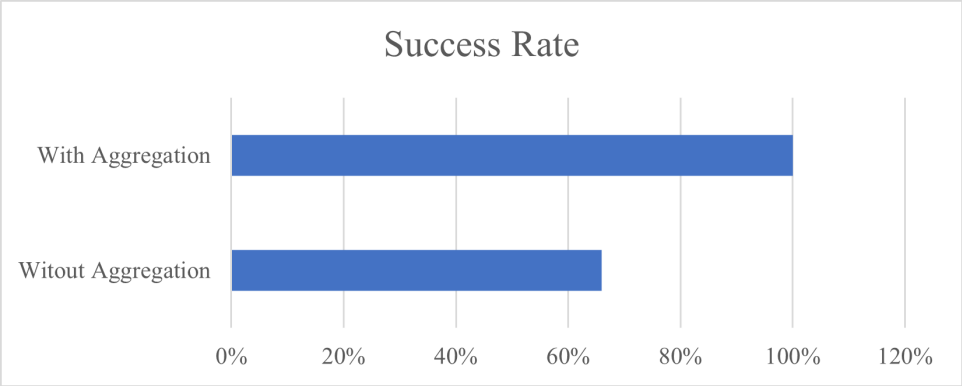
The TPS can be reduced considerably when an aggregation mechanism and a clustering method are used before sending data to the blockchain. Considering the application configurations, for 1250 mobile application users, there are only 10 TPS on the blockchain network when aggregation and clustering methods are used, significantly less than the TPS value obtained when the clustering and aggregation methods are not used. This results in lower TPS throughput and latency in the blockchain network.

Fig. 5.1 illustrates the performance comparison of the network when an aggregation mechanism is used before sending data to the blockchain and when it is not. The Figure shows that the blockchain network performs better in all the measured performance parameters when an aggregation mechanism is used.

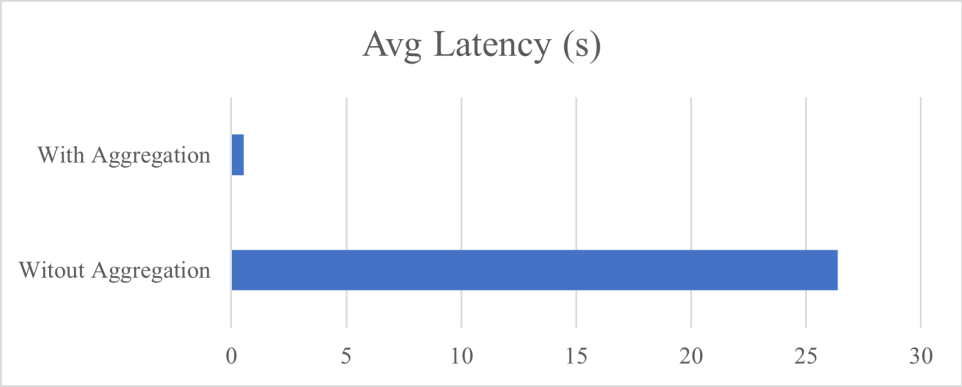
The "success rate" of transactions in the blockchain refers to the percentage of transactions that are successfully executed within the blockchain network without errors or failures. A high success rate indicates that the blockchain network can handle a significant number of transactions without errors or bottlenecks, making it more reliable. Fig. 5.1a compares success rates in two cases: when the proposed aggregation method is being used and when it is not. We can see that when the aggregation method is used, the success rate is 100%, whereas when it is not, it is much lower.

The "latency" of the network implies the time it takes for a transaction to be processed and confirmed on a blockchain network. It is a critical performance metric that measures the delay or time lag between when a transaction is submitted to the network and when it is successfully committed to the blockchain ledger. We can observe in Fig. 5.1b, that the measured average latency of the system is negligible when the aggregation method is used. This indicates that transactions are processed quickly, making the network more time-efficient and reliable.

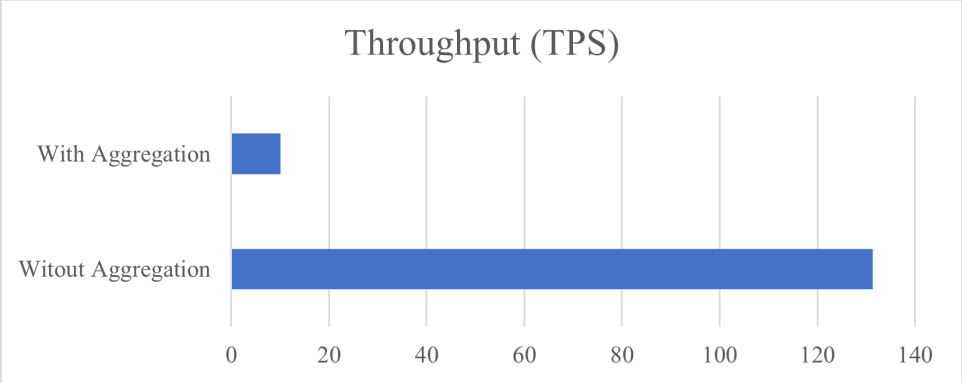
When analyzing the performance of the blockchain network, "throughput" refers to the number of transactions or operations that the network can process and confirm within a given time period. In Fig. 5.1c, the throughputs (measured in TPS) of the network in both cases are shown. With the proposed aggregation method, the TPS of the network is lower. Therefore, the throughput is maintained at a more stable value, whereas when the data is directly sent to the blockchain without aggregating, the network shows a bottleneck, and the throughput is unstable.



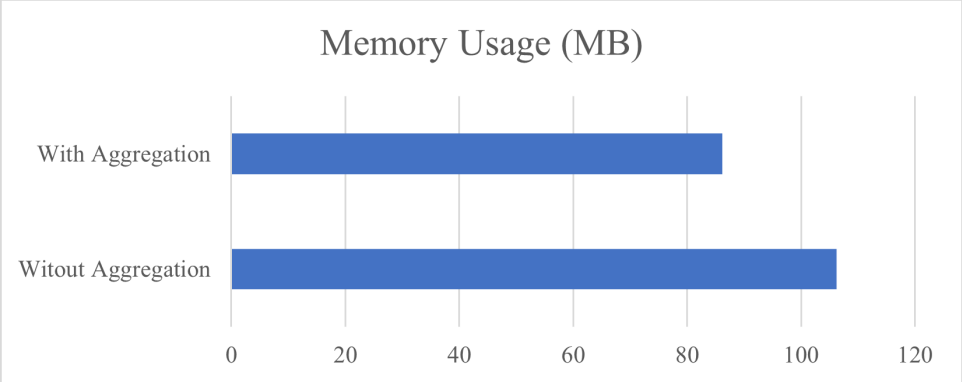
(a) Success Rate of Transactions for 10,000 Transactions.



(b) Average Latency for 10,000 Transactions.



(c) Throughput for 10,000 Transactions.



(d) Memory Usage After Running the Blockchain for 2 Minutes.

Figure 5.1: Performance Parameters of the Fabric Network.

The memory usage by the nodes in Megabytes (MB) is shown in Fig. 5.1d. The number of blocks created in a given period reduces due to low TPS. This results in less memory usage, another advantage of using the proposed aggregation mechanism.

5.2 Conclusion

This research has combined data clustering and aggregation techniques with blockchain technology to detect high RF radiation levels and enhance the system's performance. The numerical results have shown that the aggregation and clustering mechanisms have increased the performance of the blockchain network in terms of success rate, average latency, throughput, and memory usage. In addition, by aggregating the collected data, the potential faulty data levels have been eliminated as outliers, improving the overall system's reliability.

Furthermore, integrating data clustering and aggregation techniques within the Hyperledger Fabric blockchain network has yielded substantial improvements in system performance and reliability. Hyperledger Fabric, known for its robust and permissioned architecture, proved to be an ideal foundation for this research, as it facilitated the secure and efficient handling of aggregated data streams while maintaining the confidentiality and integrity of sensitive information. Additionally, the features offered in the Hypereledger Fabric blockchain framework can be used to add more functionalities to the system in the future. For example, various channel configurations in the Fabric blockchain network can extend this application's usage beyond measuring and detecting high radiation levels to store any additional private data of the service providers, which will benefit both service providers and regulators.

In conclusion, integrating data clustering and aggregation techniques with Hyperledger Fabric has led to a significantly improved blockchain system for detecting high RF radiation levels. As the technology continues to evolve, this research paves the way for innovative solutions that leverage the power of blockchain to address complex real-world challenges more efficiently and securely.

Chapter 6

Work Plan

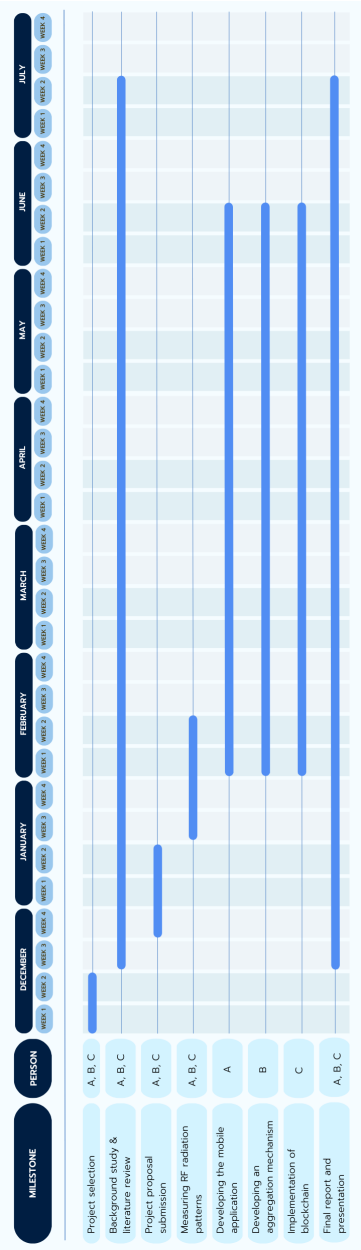


Figure 6.1: Time Plan.

Bibliography

[1] E. Commission, "Open schools journal for open science," *Journal for Open Science*, vol. 1, no. 2, pp. 320–329, 2019.

[2] L. Hardell and M. Carlberg, *Reviews on Environmental Health*, vol. 36, no. 4, pp. 585–597, 2021. DOI: doi:10.1515/reveh-2020-0168. [Online]. Available: https://doi.org/10.1515/reveh-2020-0168.

[3] B. K. Mohanta, S. S. Panda, and D. Jena, "An overview of smart contract and use cases in blockchain technology," in *2018 9th international conference on computing, communication and networking technologies (ICCCNT)*, IEEE, 2018, pp. 1–4.

[4] A. A. Monrat, O. Schelén, and K. Andersson, "A survey of blockchain from the perspectives of applications, challenges, and opportunities," *IEEE Access*, vol. 7, pp. 117 134–117 151, 2019.

[5] U. Jayarajah and A. M. Abeygunasekera, "Cancer services in sri lanka: Current status and future directions," *Journal of the Egyptian National Cancer Institute*, vol. 33, no. 1, p. 13, 2021.

[6] S. Banerjee and D. Khan, "Food for life or miasma-a study on hot food in plastic container," *International Research Journal of Modernization in Engineering Technology and Science.*, vol. 2, no. 7, 2020.

[7] T. Fiolet, B. Srour, L. Sellem, *et al.*, "Consumption of ultra-processed foods and cancer risk: Results from nutrinet-santé prospective cohort," *bmj*, vol. 360, 2018.

[8] *Electromagnetic fields and public health: Mobile phones — who.int*, https://www.who.int/news-room/fact-sheets/detail/electromagnetic-fields-and-public-health-mobile-phones, [Accessed 13-Mar-2023].

[9] J. C. Lin, "Are cell phones or radio-frequency electromagnetic fields possibly carcinogenic to humans? [telecommunications health and safety]," *IEEE Antennas and Propagation Magazine*, vol. 54, no. 1, pp. 210–212, 2012. DOI: 10.1109/MAP.2012.6202552.

[10] M. R. Miah, M. A. Hannan, A. S. Rahman, *et al.*, "Processed radio frequency towards pancreas enhancing the deadly diabetes worldwide," *Journal of Endocrinology Research*, vol. 3, no. 1, 2021.

[11] R. L. Melnick, "Commentary on the utility of the national toxicology program study on cell phone radiofrequency radiation data for assessing human health risks despite unfounded criticisms aimed at minimizing the findings of adverse health effects," *Environmental Research*, vol. 168, pp. 1–6, 2019, ISSN: 0013-9351. DOI: https://doi.org/10.1016/j.envres.2018.09.010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0013935118304973.

[12] M. Ramalingam, D. Saranya, and R. ShankarRam, "An efficient and effective blockchain-based data aggregation for voting system," in *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)*, IEEE, 2021, pp. 1–4.

[13] D. Brown, A. Japa, and Y. Shi, "A fast density-grid based clustering method," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, IEEE, 2019, pp. 0048–0054.

[14] R. Vaarandi, "A data clustering algorithm for mining patterns from event logs," in *Proceedings of the 3rd IEEE Workshop on IP Operations & Management (IPOM 2003)(IEEE Cat. No. 03EX764)*, Ieee, 2003, pp. 119–126.

[15] R. Kochhar, B. Kochar, J. Singh, and V. Juyal, *Blockchain and its impact on telecom networks*, 2018.

[16] S. Kumar and V. K. Chaurasiya, "A strategy for elimination of data redundancy in internet of things (iot) based wireless sensor network (wsn)," *IEEE Systems Journal*, vol. 13, no. 2, pp. 1650–1657, 2018.

[17] A. Ahlbom, J. Bridges, R. De Seze, *et al.*, "Possible effects of electromagnetic fields (emf) on human health–opinion of the scientific committee on emerging and newly identified health risks (scenihr)," *Toxicology*, vol. 246, no. 2-3, pp. 248–50, 2008.

[18] F. P. Miller, A. F. Vandome, and J. McBrewster, *Advanced encryption standard*. Alpha Press, 2009.

[19] E. Biham and A. Shamir, *Differential cryptanalysis of the data encryption standard*. Springer Science & Business Media, 2012.

[20] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the royal statistical society. series c (applied statistics)*, vol. 28, no. 1, pp. 100–108, 1979.

[21] K. Sasirekha and P. Baby, "Agglomerative hierarchical clustering algorithm-a," *International Journal of Scientific and Research Publications*, vol. 83, no. 3, p. 83, 2013.

[22] S. Seo, "A review and comparison of methods for detecting outliers in univariate data sets," Ph.D. dissertation, University of Pittsburgh, 2006.

[23] S. Saleem, M. Aslam, and M. R. Shaukat, "A review and empirical comparison of univariate outlier detection methods.," *Pakistan Journal of Statistics*, vol. 37, no. 4, 2021.

[24] B. Salehi, S. Jarahizadeh, and A. Sarafraz, "An improved ransac outlier rejection method for uav-derived point cloud," *Remote Sensing*, vol. 14, no. 19, p. 4917, 2022.

[25] E. Androulaki, A. Barger, V. Bortnikov, *et al.*, "Hyperledger fabric: A distributed operating system for permissioned blockchains," in *Proceedings of the thirteenth EuroSys conference*, 2018, pp. 1–15.

[26] D. Ongaro and J. Ousterhout, "The raft consensus algorithm," *Lecture Notes CS*, vol. 190, p. 2022, 2015.